DOCUMENT RESUME

ED 235 182 TM 830 468

AUTHOR	Gray, Peter J.; And Others
TITLE	A Guidebook for Conducting Field Trials of New Methods. Paper and Report Series No. 77.
INSTITUTION	Northwest Regional Educational Lab., Portland, OR. Research on Evaluation Program.
SPONS AGENCY	National Inst. of Education (ED), Washington, DC.
PUB DATE	Nov 82
CONTRACT	400-80-0105
NOTE	67p.; Not available in paper copy due to small print in Method Description Sheets.
PUB TYPE	Guides - Non-Classroom Use (055) -- Reports - Descriptive (141)
EDRS PRICE	MF01 Plus Postage. PC Not Available from EDRS.
DESCRIPTORS	*Check Lists; *Evaluation Methods; Feasibility Studies; *Field Studies; Field Tests; Formative Evaluation; Summative Evaluation

ABSTRACT

This guide provides evaluators interested in new methods with a framework for determining the feasibility, utility, and effectiveness of the methods in field settings. To accomplish this intent, the body of this guide contains discussions of five checklists. Three of the checklists are to be used before a trial to determine the feasibility of one or more methods, and during and after a trial to determine the method's utility and effectiveness. An additional checklist lists factors related to ideal field trial conditions; deviations from these ideal conditions suggest the limitations that may be imposed on a method in a particular field trial. A final checklist contains an outline for recording the conditions and results of a field test. A brief summary of 27 method trials and an updated catalog of 38 new evaluation methods are appended. (BW)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED235182

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

X This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

paper and report series

No. 77 A GUIDEBOOK FOR CONDUCTING FIELD TRIALS OF NEW METHODS

Peter J. Gray
Darrel N. Caulley
Nick L. Smith

Research on Evaluation Program

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

M. M. Rogers

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

Northwest Regional Educational Laboratory
300 S.W. Sixth Avenue
Portland, Oregon 97204
Telephone (503) 248-6800

TM 830 468

INTERIM DRAFT

Do not cite or quote
without authors'
permission.

Authors welcome reactions
and suggestions.

NO. 77 A GUIDEBOOK FOR CONDUCTING FIELD TRIALS OF NEW METHODS

PETER J. GRAY
DARREL N. CAULLEY
NICK L. SMITH

November 1982

Nick L. Smith, Director
Research on Evaluation Program
Northwest Regional Educatinal Laboratory
300 S. W. Sixth Avenue, Portland, Oregon 97204

Published by the Northwest Regional Educational Laboratory, a private nonprofit corporation. The work upon which this publication is based was performed pursuant to Contract No. 400-80-0105 of the National Institute of Education. It does not, however, necessarily reflect the views of that agency.

The information presented in this publication does not necessarily reflect the opinions of the Northwest Regional Educational Laboratory and no endorsement should be inferred.

PREFACE

The Research on Evaluation Program is a Northwest Regional Educational Laboratory project of research, development, testing, and training designed to create new evaluation methodologies for use in education. This document is one of a series of papers and reports produced by program staff, visiting scholars, adjunct scholars, and project collaborators—all members of a cooperative network of colleagues working on the development of new methodologies.

How can evaluators determine the feasibility of using a new evaluation method for a given problem? What criteria can be used in judging the utility and effectiveness of a new method in making both formative and summative assessments of the method's quality? Answers to these questions are provided in this report through a discussion of five checklists designed to aid evaluators in conducting field trials of new methods.

Nick L. Smith, Editor
Paper and Report Series

FOREWORD

The substance of this report draws upon and summarizes work completed to date by the Research on Evaluation Program (ROEP) in the field testing of alternative evaluation methods. This report, along with an earlier companion piece (Guidebook to New Evaluation Methods, Caulley and Smith, 1981, ROEP Report Number 66), is designed to aid evaluation practitioners in the selection, adaptation, and implementation of alternative evaluation approaches. These two reports will be used during the 1983 contract year in providing consultation services to evaluation practitioners. A combined, revised report will be prepared in 1984 based on this field use. The Research on Evaluation Program staff welcome reactions, comments, and revision suggestions concerning both reports.

CONTENTS

	Page
Field Trial Criteria and Checklists.	1
Feasibility and Ideal Test Conditions.	3
Formative Utility and Effectiveness.	7
Summative Utility and Effectiveness.	7
References	13
Appendix A: Summary of ROEP Method Trials	15
Appendix B: Revised Catalog of Method Descriptions. . . .	23

A GUIDE FOR CONDUCTING FIELD TRIALS
OF NEW METHODS

Field Trial Criteria and Check Lists

In recent years there has been increasing interest in the development of new methods for conducting program evaluations. This interest has been stimulated by problems of matching existing methods to diverse evaluation contexts and problems to provide results that are useful to decision makers. As a result, evaluators have explored geography, epidemiology, law, philosophy, history, art, anthropology, economics, journalism, and many other areas in search of new perspectives and tools for evaluation. (cf. Smith, 1981 (a), 1981 (b), 1981 (e))

As new approaches have been identified and described, an expanded repertoire of methods has become available for adaptation in evaluation settings. However, since few methods have yet been used for program evaluation, questions remain about their practical worth. These questions focus on the feasibility, utility, and effectiveness of the new methods.

One of the best ways to determine the worth of a method is through its implementation in a field trial. By a field trial, we mean any adaptation of a method in a field setting. Field trials may take various forms:

- In Situ Study--study of a naturally occurring use of a specific evaluation method;
- Pilot Study--field trial of a new method or of an existing method in a new context, a trial designed to study the method, not to do an evaluation;
- Application Study--conduct of an evaluation employing a new method not used before in a similar context, where the evaluation places only secondary attention on the study of the method;
- Comparative Study--an evaluation employing two or more methods, enabling comparisons to be made among them;

The overriding question which should guide a field trial is, "How well does the method work?" In terms of our three major criteria, the question is modified to be, " How well is the method likely to work?" (feasibility), "How well is it working?" (utility and effectiveness in a formative sense), and "How well did it work?" (utility and effectiveness in a summative sense). The purpose of field trials is to bring empirical, practical, and social value concerns to the fore in the development of new methods. This guide, then, provides practitioners interested in new methods with a framework for determining the feasibility, utility, and effectiveness of the methods in field settings.

To accomplish this intent, the body of this guide contains discussions of five checklists. Three of the checklists are to be used before a trial to determine the feasibility of one or more methods, and during and after a trial to determine the method's utility and effectiveness. These checklists all focus on the same elements, that is, various aspects of an evaluation such as its purpose, audience concerns, setting constraints, major questions, data collection and analysis methods, and reporting. The checklists' questions are phrased so as to explicate the criteria of feasibility, utility, and effectiveness at the three stages of trial, namely, during planning, in process, and upon completion.

An additional checklist, lists factors related to ideal field trial conditions. Deviations from these ideal conditions suggest the limitations that may be imposed on a method in a particular field trial. The limitations have to do with expectations for a successful field trial and expectations regarding the information obtained about the method's worth.

A final checklist contains an outline for recording the conditions and results of a field test.

The purpose of this guide is to provide practitioners with practical method review guidelines, therefoe, more technical discussions of the development and testing of methods have not been included here. (See, for example, Smith 1979, Smith 1981 (a), Smith 1981 (e), and Smith 1982 (a).) A brief summary of

the 27 methods trials performed to date by the Research on Evaluation Program (ROEP) has been included as Appendix A for the reader's future reference. A wide variety of method trials are presented in Appendix A, providing the reader with a range of alternative test strategies to consider.

Appendix B of the guide contains on updated catalog of new evaluation methods. Thirty-eight new methods are described in terms of their purpose, basic procedures, advantages/benefits, disadvantages/costs, and resources required. The basic references provided for each method enable the reader to learn more about particular methods of interest.

Feasibility and Ideal Test Conditions

The more new methods there are, the more difficult it becomes to decide which one(s) to adapt for a field trial. Questions arise concerning the potential worth or feasibility of a method within the constraints of an evaluation's context.

The feasibility checklist (Checklist 1) contained on the next page can be used with the catalog of new methods (Appendix B). For example, to determine if a method is consistent with the purpose(s) of the evaluation (Item 1), one would review the method's purpose. The information related to why and when to use a method is useful in answering Item 2, will the method probably provide results that are appreciated by major interest groups?

Reviewing the basic procedures of a method will help a potential user decide if the evaluator(s) can learn the method sufficiently to conduct a credible evaluation (Item 3). Reviewing this part of the method description will also reveal the extent to which the method has been developed. The basic references section provides additional information in this regard.

CHECKLIST 1: FEASIBILITY

1. Is the method consistent with the purpose(s) of the evaluation?

2. Does the method promise to answer the evaluation questions posed by the major interest groups?

3. Can the evaluator(s) learn the method sufficiently well to conduct a credible evaluation?

4. Is it possible to adapt the method within the constraints of the evaluation setting and still use it validly?

5. Is there a need for this method, that is, does it provide a unique approach, or are there better methods?

It is very difficult to decide before the fact if it is possible to adapt a method within the constraints of the evaluation setting and still use it validly (Item 4). However, the list of factors that provide for ideal test conditions (included here as Checklist 2) provide a framework for analyzing the setting in order to make that decision. The first five factors concern the context of the evaluation, that is, agency support, program clarity, conflicting influences and so on.

The evaluator's capabilities are assessed relative to the method in factor 6, and the extent of the method's development is considered in factor 7.

Factors 8 and 9 focus on the interaction of the method and the program; the match between them in a general sense (8) and the availability of sufficient resources to implement the method (9). A careful reading of the why and when to use, advantages/benefits and disadvantages/costs sections of the method description can give the evaluator an idea of the likely match between the evaluation context and the method. From these sections and the resources required section, come an understanding of the resources needed to adequately implement the method.

The information gathered relative to these factors will indicate the degree to which trial conditions vary from the ideal. It will be up to the evaluator to determine if the deviation is sufficient to decide against trying the method being considered. The final item on the feasibility checklist, regarding the uniqueness of the method, requires that the evaluator look not only at new methods but at existing methods as well. In the final analysis, it may simply be unnecessary and unwise to try a new method when a familiar method will do.

CHECKLIST 2:
FACTORS THAT PROVIDE FOR IDEAL TRIAL CONDITIONS

1. The agency sponsoring the evaluation should be supportive of trying the new approach.

2. The project or program being evaluated should be clearly defined and fully implemented.

3. There should be no conflicting political, social, or organizational influences shaping the evaluation.

4. The setting should allow maximum flexibility for adapting the method to local conditions.

5. There should be internal support by the program staff or product developers for the evaluation of the new method.

6. The test should be conducted by seasoned practitioners familiar with the new method and with the evaluation of similar programs or products in such settings.

7. The method should be as completely developed as possible prior to the empirical test.

8. There should be an optimum match between the method, the program or product being evaluated, and the evaluation context.

9. There should be ample resources (e.g., time, money, people, materials) to conduct the evaluation using the new method.

Formative Utility and Effectiveness

Once a method has been chosen, its use in a field trial must be monitored to determine how it is performing and to determine what modifications need to be made to make it effective. The items on the formative utility and effectiveness checklist (Checklist 3) are based on the standards associated with utility developed by the Joint Committee (1981). They concern in order: evaluator credibility, information scope and selection, audience identification, resource utilization, and evaluation impact. The last item provides a check on the fidelity with which the method is being used.

Checklist 3 may be used by the evaluation staff or by others to conduct a formative meta-evaluation of the method. It can be helpful to have someone not directly involved in the use of the method act as a critic since, in their enthusiasm for using the method, the evaluators may lose sight of the larger picture represented by the utility and effectiveness checklist items.

Summative Utility and Effectiveness

At the conclusion of a field trial it is appropriate to take a retrospective look at the method and its implementation. The summative utility and effectiveness checklist (Checklist 4) has items on it which, when answered, generate the information to be included on the Field Trial Report form (Checklist 5). There are six categories of information on each of these two instruments.

Ideal conditions factors. Before the field trial, factors that provide ideal conditions were considered in order to help determine the feasibility of trying a particular method. Reviewing the factors again at the end of a trial is a good starting point for assessing the method's worth since the deviations from the ideal indicate the context in which the trial took place.

CHECKLIST 3: FORMATIVE UTILITY AND EFFECTIVENESS

1. Is the method being used by the evaluator(s) in a competent way?

2. Does the method facilitate the collection of information sufficient to address the pertinent questions of the evaluation?

3. Does the implementation of the method provide initial findings that are understandable by the major interest groups?

4. Are the resources being used by the method consistent with initial projections:

   - time
   - cost
   - people
   - materials

5. What are the unanticipated outcomes or side effects of implementing the method?

6. Is the implementation of the method remaining true to its basic assumptions or principles?

CHECKLIST 4: SUMMATIVE UTILITY AND EFFECTIVENESS

1. What deviations from ideal conditions were present in the trial?

2. a. Did the method serve the desired purpose?

   b. Were the important evaluation questions answered by the method?

3. a. Did the data collection and analysis produce technically accurate results?

   b. Were the collected results analyzed and summarized in a way appropriate to the problem?

   c. Were the results of the evaluation useful to the major interest groups?

4. Did the method produce an evaluation that met the resource conditions of

   - time
   - cost
   - people
   - materials

5. a. Was the method relatively easy to learn by the evaluator(s)?

   b. Was the method sufficiently adaptable to be rated effective in this evaluation setting?

   c. What were the method's major strengths and weaknesses in this context?

   d. What were the unanticipated outcomes or side effects of using the method?

   e. Are there recommendations that should be passed on to other potential users?

CHECKLIST 5: NEW METHOD FIELD TRIAL REPORT OUTLINE

Method Name

Program/Product (brief description)

Trial Conditions Profile

- Program factors
- Evaluator capabilities
- Method development
- Method/Program match
  - purpose
  - procedures
- Resource Utilization
  - time
  - money
  - people
  - materials

Results' Quality

- Technical accuracy of data collection and analysis
- Appropriateness to the problem
- Usefulness to major interest groups

Method Analysis

- Ease of learning and use
- Adaptability
- Strengths/weaknesses
- Unanticipated outcomes or side effects
- Recommendations for future users

Field Trial Contact Person

Other Contact People or Resources

Program related elements constitute the first five factors. They concern such things as agency and internal support, program clarity, conflicting influences, and flexibility. There may well be a change in one's analysis of these factors after a trial has been run. Resistance may have developed, conflicts may have arisen, and old prejudices for particular methods may have reduced the amount of flexibility available for adapting the methods.

By referring back to the factors, an evaluator can generate a test's profile as shown on the Field Trial Report (Checklist 5). Under the method/program match, the evaluator is encouraged to give a brief summary of the purpose and procedures of adapting the method within the given setting. The use of resources is summarized under the last factor.

Results' quality. The quality of information resulting from the use of a method is an indication of the method's utility and effectiveness within a given setting. By looking at the results, one may infer the worth of the method which produced them. There are three things to consider in assessing the quality of results. The following concepts of quality are based on the accuracy standards of the Joint Committee (1981).

The first aspect of quality has to do with the technical accuracy of the data. Of concern here is the validity and reliability of data collection, the system for managing data quality control, and the appropriateness of the analysis procedures to the evaluation problem.

The second aspect of quality is the appropriateness to the problem of the information collected, analyzed, and summarized. The basis for determining appropriateness lies in looking at the data within the context of the evaluation. The description of the trial conditions provides this perspective.

Finally, the usefulness of the results to major interest groups must be considered in determining their quality. In this regard, conclusions must be justified in the eyes of decision makers and the reporting of results must be viewed as being objective.

Method analysis. The final set of checklist items and the final set of report topics focus on the more complete description of the method. Included here are the ease of learning and using the method, its adaptability, strengths and weaknesses, and any unanticipated outcomes and side effects resulting from the use of the method.

Information from this last part of the checklist and report should help a prospective evaluator make a final decision about trying a method. For example, the ease with which a method is learned and used, together with the information about the evaluator's capabilities presented earlier in the report, should help people judge if the method is right for them. In addition, an assessment of the extent to which a method can be adjusted in the process of making it effective will provide insight as to its robustness. The potential user can refer to this information after analyzing his or her capabilities, the evaluation setting, and the "strain" that these will put on the method as part of a feasibility check.

After noting strengths and weaknesses of the method, unanticipated outcomes and side effects and any recommendations for future users, an evaluator can record more personal insights and anecdotes. These give the flavor of a trial and are often the most interesting parts of a report.

The name and address of the field trial contact and any other resources that the evaluator has become aware of as part of the field trial are included at the end of the Field Trial Report.

Until there is concrete information about the feasibility, utility, and effectiveness of new methods, evaluation practitioners must make choices based on superficial characteristics and projected benefits. What is needed is a body of literature describing in a consistent way information, from field trial experiences, about the worth of methods. It is hoped that the checklists provided here will help practitioners select, implement and evaluate the quality of new methods as well as provide documentation about their experiences that can be shared with other practitioners.

REFERENCES

Joint Committee. Standards for the evaluation of educational programs, products and materials. New York: McGraw Hill, 1981.

Levin, H.M. Cost-Effectiveness: A Primer. Beverly Hills, CA: SAGE Publications, 1983.

Smith, N.L. Classic 1960s articles in education evaluation. Evaluation and Program Planning, 1981, 4 (2), 177-182. (c)

Smith, N.L. (Ed.) Communication strategies in evaluation. Beverly Hills, CA: SAGE Publications, 1982. (b)

Smith, N.L. Evaluating evaluation methods. Studies in Educational Evaluation, 1981, 7(2), 173-181. (d)

Smith, N.L. (Ed.) Federal efforts to develop new evaluation methods. San Francisco, CA: Jossey Bass, 1981. (e)

Smith, N.L. (Ed.) Field assessments of innovative evaluation methods. San Francisco, CA: Jossey Bass, 1982. (a)

Smith, N.L. (Ed.) Metaphors for evaluation: Sources of new methods. Beverly Hills, CA: SAGE Publication, 1981. (a)

Smith, N.L. (Ed.) New techniques for evaluation. Beverly Hills, CA: SAGE Publications, 1981. (b)

Smith, N.L. Requirements for a discipline of evaluation. Studies in Educational Evaluation, 5, 1979, 5-12.

Smith, N.L. and Caulley, D.N. The evaluation of educational journals through the study of citations. Educational Researcher, 1981, 10 (5), 11-12, 22-24.

APPENDIX A

Summary of ROEP Method Trials

1. N. Stenzel Committee Hearings as an Evaluation Format
(Available in Smith, 1982 (a))

What is the nature of committee hearings as used in congressional inquiry and policy formation? Can committee hearings be adapted for use in educational evaluation? This report discusses these and other questions in assessing the potential of committee hearings as a tool in evaluation. Included in this report is a sample application of a committee hearing procedure in evaluation and an extensive discussion of the various elements of the committee hearing process, including committee operation, staffing, and procedural rules.

2. P. A. Templin Photography as an Evaluation Technique
(Available in Smith, 1982 (b))
(Prior version available as ROEP Report #32)

Probably the most difficult aspect of developing new approaches to evaluation is in conducting the first meaningful applications of them in field settings. This document contains a report of a field trial of the use of photography as an evaluation method. A photographic study of a professional conference was conducted. Although the study was accomplished under severe time restrictions, this report illustrates the considerable promise of photographic evaluations. This document, written to help others who wish to conduct photographic studies, includes discussions of the working theories that can be used to guide the photo taking, the technical details and practical problems of concern in such a study, as well as the results of this particular conference photo study.

3. N. L. Smith Meta-Evaluation: Alternative Perspectives
(Available in Smith 1981 (b))

What range of approaches are available for the evaluation of evaluations (meta-evaluation)? Will differing approaches to meta-evaluation result in differing assessments of an evaluation's worthiness? These and related questions are addressed in this report which presents four comparative evaluations of the same report: "Some Still do: River Acres, Texas" by Terry Denny. The concept of meta-evaluation is significantly broadened by these meta-evaluations which use watercolor painting as a perspective (William J. Gephart), film criticism as a perspective (Gabriel Della-Piana), philosophic analysis as a perspective (D. Bob Gowin), and product evaluation as a perspective (Jason Millman). These are stimulating alternatives to consider as the need for quality meta-evaluations becomes more prevalent.

4. N. L. Smith — Investigative Tracking as a Method of Library Evaluation
(Available in Smith, 1982 (a))

Journalists have developed a simple strategy for using archival records in conducting evaluative investigations. This strategy provides a straightforward, cost-effective means of evaluating library operations and is especially suited to informal studies of small special libraries.

5. D. E. Nelson — Investigative Journalism Methods in Educational Evaluation.
(Available in Smith, 1982 (a))

How might one apply the methods of investigative journalism in evaluation studies performed in state education agencies? How would such methods compare with the more traditional approaches to evaluation? Dr. David Nelson of the Utah State Office of Education addresses these concerns by presenting two comparative cases which illustrate the strengths and weaknesses of investigative journalism methods in state department settings. Dr. Nelson uses the newly developed Standards for Educational Evaluation as a yardstick for assessing the relative merits of this new approach over conventional methods.

6. B. McGaw — The Use of Exploratory Data Analysis in Educational Analysis
(Available in Smith, 1981 (b))

What is the nature of exploratory data analysis, and how does one perform such analysis? How might these techniques be used in educational evaluation? Barry McGaw considers these questions in this paper. Using data from evaluation studies in education, Dr. McGaw leads the reader through a step-by-step summary of the basic techniques employed in exploratory data analysis. With the aid of this paper, evaluators can begin to shift their perspective from confirming assumed relationships to exploring possible relationships in educational programs.

7. N. L. Smith and D.N. Caulley The Evaluation of Educational Journals Through the Study of Citations. (Available in Smith and Caulley, 1981)

A citation study of all major articles published in AERJ, RER, and ER in 1972 and 1973 was conducted to evaluate the impact of these journals on the educational literature. RER was found to be making the most significant impact in terms of subsequent citations of articles. While the analysis revealed a relatively closed network of interjournal communication within education for these articles, no discernible yearly citation pattern was identified for the individual articles.

8. N. L. Smith Classic 1960s Articles in Educational Evaluation (a citation analysis application) (Available in Smith, 1981 (c))

Which of the early 1960s literature in educational evaluation is still being widely read today? A citation study was conducted to identify which articles on evaluation theory and methods written in the 1960s continued to be heavily read and cited through the 1970s. While most of the writing of the 1960s has ceased to have wide appeal, five classic articles were identified which continue to significantly influence thinking in evaluation.

9. C. A. Lane Using the Tools of Philosophy: Metaphor in Action (Available in Smith, 1982 (a))

How might the methods of philosophy be applied to the evaluation of educational programs? In what ways would evaluations using philosophical concept analysis look different from traditional studies? This paper presents a retrospective analysis of the possible use of philosophic techniques in a major reading evaluation conducted in a middle-sized school district. By working with the district evaluation personnel, the author of this report explores how that evaluation might have been different if the evaluation staff had employed philosophic procedures.

10. J. Stanfield Management Review of Evaluation Practice (Available as ROEP Report #58)

What is the nature of evaluation practice in state departments of education when viewed by a management consultant rather than a traditional educational researcher? What management perspectives might be used to better understand the operation of state department evaluation units? These and related

questions are answered in this report by Dr. Jonathan Stanfield, an independent management consultant who conducted a review of the state evaluation units in Washington, California, and Montana. This report provides an illustration of a management consulting review study in evaluation as well as offering insightful observations about the nature of evaluation practice in state departments.

11. J. Stanfield Pilot Field study of SEA Evaluation Costs
(Available as ROEP Report #69)

What are the costs associated with evaluation at the state level? Can management consulting techniques be used to be used to reduce some of these costs? After summarizing the costs of evaluation today (approximately $32 million at the state level), Dr. Stanfield reviews the ways in which management consulting techniques could replace standard evaluation approaches or save as much as $2.9 million. Important distinctions in considering evaluation costs are also included in this pilot study report.

12. J. Stanfield Management Consulting Case Study
(Available as ROEP Report #74)

Can management consulting techniques provide cost effective alternatives to standard evaluation approaches? What is the cost effectiveness of using management consulting techniques in SEA evaluation unit settings? This field study addresses these questions and concludes that management consulting is a competitive approach, but due to the limited applicability of management consulting to the study exercises selected and the lack of controls for comparison, it is not possible to show conclusively that management consulting is either superior or inferior to standard evaluation approaches.

13. H. M. Levin and W. Seidman Manual for Cost Analysis in Educational Evaluation
(Available in Levin, 1983)

How does one select the appropriate cost analysis methods for use in an evaluation study? What are the procedures for estimating costs and benefits and for combining them to provide suitable cost analysis information? These and related issues are discussed in this report which is an extension of an earlier document (ROEP Report #33, Case Book on Cost Analysis in Educational Evaluation). This manual provides

training in performing cost-benefit, cost effectiveness, cost utility, and cost feasibility analysis and includes major sections on establishing a decision framework; identifying, measuring, and distributing costs; and assessing and aggregating benefits, effects, and utilities.

14. G. M. Della-Piana Film Criticism and Micro-Computer Couseware Evaluation
(Available in Smith 1982 (a))

How might techniques from film criticism be used in the evaluation of instructional materials? Based on previous project work on methods of criticism, this report suggests an innovative system for providing in-depth reviews of micro-computer-based instructional packages. This system, which emphasizes the illumination of curricular data relations through engaging criticism techniques is designed to help audiences see how a piece of courseware works and to involve audiences in making their own interpretations and judgments of it.

15. G. M. Della-Piana and C. K. Della-Piana Making Courseware Transparent: Beyond Initial Screening
(Available as ROEP Report #76)

How can one evaluate microcomputer courseware in a way that helps others make their own judgments of its quality? What alternative styles of evaluative presentations are available? This report answers these and related questions by describing a collection of procedures, with illustrative examples, for selecting and portraying evaluative information. A checklist for conducting microcomputer courseware criticism is also included.

16. B. E. Wholeben and J. M. Sullivan Multiple Alternatives Modeling in Determining Fiscal Roll-Backs During Educational Funding Crises
(Available as ROEP Report #70).

How can one determine the proper mix of educational programs to receive reduced funding when budget cutbacks are necessary? This report provides an extensive discussion of the use of criterion referenced, mathematical modeling procedures to determine which budget reductions minimally reduce the quality of educational programs. Part I of this report explains the basic design of multiple alternatives analysis and the context for its use; Part II provides the technical and mathematical details of the analysis; and Part III contains an

extensive example of the use of these procedures in reducing a budget within a local school district. This report describes a highly technical but workable solution to the difficult problem of reducing school budgets.

17. A. Rasp, Jr. Interviewing to Augment Large Scale Survey Data: The Washington High School and Beyond Story
(Available on ROEP Report #71)

Is interviewing using trained volunteers a cost-effective way of providing supplemental data to increase the local relevance of national survey information? Can small-scale studies be effectively piggy-backed onto larger national efforts to increase their local utility? Based on the project described here, the answer to both these questions is yes. This project, mounted by the testing and evaluation unit of the Washington state department of education, illustrates a cost-effective way to gather supplemental qualitative and quantitative information of increased local utility by properly augmenting a national study--illustrating an important design option in this period of decreasing evaluation resources.

Applications of Alternative Evaluation Communication Techniques
(Available in Smith 1982(b))

18. D. J. Macy Research Briefs.

19. W. E. Hathaway Graphic Display Procedures.

20. G. Leinhardt and S. Leinhardt Stem-and-Leaf Displays.

21. B. E. Wholeben Operational Network Displays.

22. N. L. Smith Geographic Displays.

23. M. Hendricks Oral Policy Briefings.

24. R. E. Stake and D. E. Balk Briefing Panel Presentations.

25. M. Levine Adversary Hearings.

26. N. Stenzel Committee Hearings.

27. J. S. Shoemaker Television Presentations.

APPENDIX B

Revised Catalog of Method Descriptions

Revised Catalog of Method Descriptions

		Page
Data Collection and Analysis Strategies		
1.	Assignment Models.	27
2.	Transportation Models.	28
3.	Dynamic Programming.	29
4.	Queueing Theory.	30
5.	Minimum-Maximum Goal Projection.	31
6.	Geocode Analysis	32
7.	Trend Surface Analysis	33
8.	Social Area Analysis/Ecological Analysis	34
9.	Concept Analysis	35
10.	Thematic Matrix Analysis	36
11.	Document Analysis.	37
12.	Legislative History.	38
13.	The Key Interview.	39
14.	Interviewing: Circling, Shuffling and Filling .	40
15.	Documenting Files and Summaries.	41
16.	Cost-Benefit Analysis.	42
17.	Cost-Effectiveness Analysis.	43
18.	Cost-Utility Analysis.	44
19.	Cost-Feasibility Analysis.	45
Sampling Strategies		
20.	Blanket Sampling	46
21.	Shadow Sampling.	47
22.	Time-Based Sampling.	48
23.	Event-Based Sampling	49
24.	Dimensionally Based Sampling	50
Reporting Strategies		
25.	Research Briefs.	51
26.	Appeals Procedures	52
27.	Storytelling	53
28.	Compelling the Eye	54
29.	Representation of Reality.	55
30.	Accurate, Sharp Descriptions	56
31.	Graphic Displays	57
32.	Stem and Leaf Displays and Box Plots	58
33.	Still Photography.	59
34.	Oral Briefings	60
35.	Briefing Panel Presentations	61
36.	Adversary Hearings	62
37.	Committee Hearings	63
38.	Television Presentations of Hearings	64

Method Description Sheet 1

Method: Assignment Models from Operations Research

Purpose: To assign something to something so some minimum or maximum is achieved.

Why and When to Use: Examples of when to use the assignment model are as follows: Tutors are to be assigned to students so that tutors' preferences are either minimized or maximized. Teachers are to be assigned to courses so that the teachers' preferences are either maximized or minimized. We have buses at various locations and we want to minimize the miles traveled to pick up pupils at various sites.

Basic Procedures: A simple algorithm is available to solve the problem by hand.

Advantages/Benefits: An advantage of the assignment model is that it gives a better solution than can be obtained by inspection. The algorithm given for solution of the problems can readily be solved by hand without the problem of gaining access to a computer. A solution is available in a very short time.

Disadvantages/Costs: The algorithm is a little tricky to learn.

Resources Required: None.

Basic References:

Caulley, D. N. The use of assignment and transportation models in evaluation. No. 68 of the Paper and Report Series of the Research on Evaluation Program. Portland, OR: Northwest Regional Educational Laboratory, 1981.

Eck, R. D. Operations research for business. Chapter 8. Belmont, CA: Wadsworth, 1976.

Hillier, F. S., & Lieberman, G. J. Introduction to operations research. Chapter 6. San Francisco: Holden-Day, 1967.

Trueman, R. E. An introduction to quantitative methods in decision making. Chapter 8. New York: Holt, Rinehart and Winston, 1977.

Method: Transportation Models

Purpose: To assign teachers to courses when courses each have a number of sections and teachers each have a number of class periods available. Note that this is different from the assignment model where each course has one section and each teacher has one class period available.

Why and When to Use: Suppose there are a number of courses to be taught and a number of teachers to do the teaching. These two numbers need not be equal. The teachers have available various numbers of class periods. There are a number of sections for each course that are required to be taught. In order to apply the model, the total number of class periods available must equal the total number of sections that are required to be taught. For each course the teachers have given a preference rating according to whether they would like to teach a course. For example, a seven-point rating might be used, with 1 representing high preference and 7 representing low preference. The problem is to assign teachers to sections and courses so as to minimize the sum of the ratings.

Basic Procedures: An algorithm is available to solve the problem by hand.

Advantages/Benefits: An advantage of the transportation model is that it gives a better solution than can be obtained by inspection. The algorithm given for solution of the problems can readily be solved by hand without the problem of gaining access to a computer. A solution is available in a very short time.

Disadvantages/Costs: The algorithm is a little tricky to learn.

Resources Required: None.

Basic References:

Caulley, D. N. The use of assignment and transportation models in evaluation. No. 68 of the Paper and Report Series of the Research on Evaluation Program. Portland, OR: Northwest Regional Educational Laboratory, 1981.

Eck, R. D. Operations research for business. Chapter 8. Belmont, CA: Wadsworth, 1976.

Hillier, F. S., & Lieberman, G. J. Introduction to operations research. Chapter 6. San Francisco: Holden-Day, 1967.

Trueman, R. E. An introduction to quantitative methods in decision making. Chapter 8. New York: Holt, Rinehart and Winston, 1977.

Method: Dynamic Programming

Purpose: To help a decision maker make a multi-stage decision, one that requires several interrelated choices.

Why and When to Use: Dynamic programming is useful in curriculum design. Suppose we have a mass of materials or topics, all are possibilities for inclusion in a course of study, but we are limited in time. How do we cull out what is unnecessary? How do we give the remaining material its proper weight, so that all will fit within the total time allotted, and achieve the maximum possible benefit? This is a familiar problem, and a subjective one; examples are: a university unable to decide what should be contained in a "liberal education;" a teacher of U. S. history somehow unable to teach past the Civil War; and a district curriculum committee debating over which readings should be included in tenth grade literature.

Basic Procedures: There are at least five requirements which appear necessary to solve the "curricular problem."

1. A list of curriculum materials. This list might be of classes which could be desirable for a training sequence; reading assignments for a history course; homework problems for general mathematics; field trips or films for science.

2. Values for each topic. These would be subjective ratings, such as might be assigned by an experienced teacher.

3. Costs for each topic. These may be expressed in terms of any resource which is limited. In curricular cases, the most usual constraint is student time.

4. Maximum cost permitted for all topics. Usually there is only a limited amount of time available for the total intended course of study.

5. A solution system. To select a final set of topics from the cand'date list (1 above), we need some way to combine the values of each topic (2) and its cost (3), such that the sum of all such costs for the chosen topics will not exceed the maximum permitted (4), but such that, at the same time, the sum of their values will be the highest possible for the data used. Algorithms are available to solve the dynamic programming problem.

Advantages/Benefits: The human judge will ordinarily not do as well as the application of dynamic programming.

Disadvantages/Costs: It is difficult to get the dynamic programming model to mirror real life examples.

Resources Requires: None.

Basic References:

Page, E. B., Jarjoura, D., & Konopka, C. D. Curriculum design through operations research. American Educational Research Journal, 1976, 13, 31-49.

Eck, R. D. Operations research for business. Belmont, CA: Wadsworth, 1976.

Hillier, F. S., & Lieberman, G. I. Introduction to operations research. San Francisco: Holden-Day, 1967.

Trueman, R. E. An introduction to quantitative methods in decision making. New York: Holt, Rinehart and Winston, 1977.

Method Description Sheet 4

Method: Queueing Theory

Purpose: Queueing Theory can be used to study waiting line problems

Why and When to Use: Queueing Theory is used whenever there are delays in waiting lines. For example, students often wait in line at libraries, cafeterias, and showers. They also wait in line at the offices of school nurses, vice-principals, and counselors. One of two errors will often be present: Either customers are waiting for service, or servers are waiting for customers. Most questions will center around three major questions: (1) How much of the time will the service channels (e.g. library checkout desk) be idle? (2) How many customers (e.g. students) will usually be waiting for service? (3) How long of a wait will each customer typically have?

Basic Procedures: Mathematical models are fitted to aspects of queueing behavior. If arrivals are assumed to be random, the probability that there will be no student arriving at the service channel is an exponential function of the number of minutes elapsed since the last arrival. One key distribution that follows from this is called the Poisson distribution. From this information it is possible to derive formulas which give the probability of finding a service channel idle at any particular time, the probability of having a certain number of customers either being served or waiting in line, the mean queue length, the expected number of customers waiting for service, the expected waiting time of a new arrival, and the probability of having a waiting line longer than a certain number of customers .

Advantages/Benefits: Using the measured information of the number of students arriving per minute, and the number of students serviced in a minute, it is possible to calculate a wide array of information about the queueing system which is useful for making decisions about it.

Disadvantages/Costs: Within queueing theory there is no automatic answer. The theory does not lay out decision alternatives. For example, in the case of the library checkout desk, we might calculate the waiting time and decide that it is unacceptably large. Note that queueing theory does not help us make the decision of whether or not the waiting time is unacceptably long.

Resources Required: None.

Basic References:

Eck, R. D. Operations research for business. Chapter 8. Belmont, CA: Wadsworth, 1976.

Hillier, F. S., & Lieberman, G. J. Introduction to operations research. Chapter 6. San Francisco: Holden-Day, 1967.

Page, E. B. Educational evaluation through operations research. No. 30 of the Paper and Report Series of the Research on Evaluation Program. Portland, OR: Northwest Regional Educational Laboratory, 1979.

Trueman, R. E. An introduction to quantitative methods in decision making. Chapter 8. New York: Holt, Rinehart and Winston, 1977.

Method Description Sheet 5

Method: Minimum-Maximum Goal Projection

Purpose: The purpose is to show the client the minimum and maximum that the evaluator can be expected to achieve.

Why and When to Use: In the process of planning an evaluation, the evaluator does a preliminary study in order to determine the feasibility of the evaluation. The question is whether the results to be expected from the evaluation will repay the resources that will have to be expended on it. Evaluators specify modest minimum expectations which they were *certain* they could meet and then specify other expectations which they *might* be able to meet if everything fell into place (e.g., sampling could be accomplished in accordance with sampling principles; political intervention could be eliminated). These minimum and maximum projections may be presented to the client. This would make it very clear to the client what he/she can expect from the evaluation and what he/she is getting for his/her money. The client *may* decide to increase the funds available to the evaluator so that the minimum goals may be increased. On the other hand, the client might decide that he/she is not getting enough from the evaluation and thus negotiate with the evaluator the raising of the minimum projection.

Basic Procedures: In planning an evaluation, the evaluator makes projections of the minimum and maximum that can be achieved.

Advantages/Benefits: The evaluator and the client should have clear expectations of what can be expected from the evaluation. The evaluator cannot be accused of over-promises.

Disadvantages/Costs: The time and cost in carrying out the planning of an evaluation will be increased. It is difficult to make realistic projections of the minimum and maximum goals to be achieved.

Resources Required: Extra time and cost in planning an evaluation.

Basic References:

Guba, E. G. Investigative journalism. In N. L. Smith (Ed.), *New techniques for evaluation*. Beverly Hills, CA: SAGE Publications, 1981.

Williams, P. N. *Investigative reporting and editing*. Englewood Cliffs, NJ: Prentice-Hall, 1978.

Method: Geocode Analysis

Purpose: Geocode analysis is a technique developed in the field of geography for displaying and analyzing geographically related information.

Why and When to Use: Many evaluation studies concern the analysis of geographic problems: the identification of individuals needing specialized attention in health, welfare, and education; the busing of children to achieve equal educational opportunity. Geocode analysis has been used in studies of school redistricting, the identification of Title I students, and in projections of student population growth. In evaluation, the technique has been used to plot differential reading growth due to remedial interventions at selected sites within a district. The technique is especially recommended for settings such as large school districts, community health service centers, and local law enforcement agencies which have pre-existing data files and geographically defined service districts.

Basic Procedures: Geocode analysis uses the individual as the unit of analysis and aggregrates individual data over geographic areas. Following are the steps involved: (1.) compile an Address Coding Guide (e.g., address of each student); (2.) builds an Individual Characteristics Data File (e.g., student achievement scores); (3.) merge the Address Coding Guide with the Individual Characteristics Data File; and (4.) produce computer grids, plots, and contour maps of individual characteristics by geographic location.

Advantages/Benefits: With geocode analysis one can examine geographically related data on thousands of individual cases. Computer-generated maps can be produced at various scales (several blocks, an entire school district, contiguous service areas, etc.) and at multiple time periods.

Disadvantages/Costs: Geocode analysis requires considerable manpower, money, and computer resources to develop and maintain an address coding guide, to merge the guide with existing individual records systems, and to produce computer-generated grids and maps.

Resources Required: Access to a computer and a computer program. Clerical help is needed to compile an Address Coding Guide and an Individual Characteristics Data File.

Basic References: A detailed description of developing and maintaining digitized address coding guides, individual characteristics data files, and geographic data bases can be found in:

McIsaac, D. N., Spuck, D. W., & Stofflet, F. P. A users guide to a system of programs for the analysis of geographic areas. Madison, WI: Information Systems for Education, 1972.

Spuck, D. W. Data base considerations. Paper presented at the Annual Meeting of the American Educational Research Association, Chicago, Illinois, 1972. (ERIC Number ED 064 894)

For indepth discussions of geocode analysis and examples of applications, the reader is referred to:

Costa, C. H. Application of geocode and mapping. Paper presented at the Annual Meeting of the American Educational Research Association, Chicago, Illinois, 1972. (ERIC Number ED 062 765)

Spuck, D. W. Geocode analysis. In H. J. Walberg (Ed.) Evaluating educational performance. Berkeley: McCutchan, 1978.

For a general introduction to geocode analysis, see:

Smith, N. L. Techniques for the analysis of geographic data in evaluation. Evaluation and Program Planning, 1979, 2, 119-126.

Method: Trend Surface Analysis

Purpose: Trend surface analysis is a technique from geology used to generate three-dimensional contour maps on which to illustrate changes in important variables. Variations in a particular variable over a geographic regions are partitioned into broad regional trends and small-scale local deviations from these trends.

Why and When to Use: Trend surface analysis has been used to study the relationship between geographic location and the following: school board elections and sources of local school support; the dissemination of information on Title III projects; statewide educational needs assessment; factors such as personal income, educational background, unemployment, number of dependents and financial support for local education.

Basic Procedures: 1. Stations are the points in the geographic region chosen to represent a local geographic area. For example, in a study of fourth grade IQ scores, each school would be a station and the average IQ at the fourth grade level would be the station value. 2. Each station is identified by three data points: x, y coordinates which establish the station's geographic location, and the station value for the variable of interest, the z coordinate. A statistical modeling procedure is employed (e.g., least squares polynomical trend fittings) and a surface equation is produced. 3. The surface equation can then be used to construct a contour map indicating regional trends and local trends. 4. Residuals can be computed and residual maps plotted to identify hidden trends and deviant stations. 5. Multiple surface maps can be overlayed to illustrate regional interactions between variables of interest.

Advantages/Benefits: Trend surface analysis is a useful technique for the visual display of large amounts of data. Using computers and high-speed plotters, the evaluator can generate multiple data maps which portray geographic data relationships not discernible through tabular display.

Disadvantages/Costs: The technique requires computer analysis and plotting facilities.

Resources Required: A computer program and access to a computer with plotting facilities.

Basic References: The following provide general introduction to trend surface analysis:

Lewis, M. S. Trend surface analysis of community variables. Psychological Bulletin, 1977, 84, 940-949.

McIsaac, D. N. The application of trend surface analysis. Paper presented at the Annual Meeting of the American Educational Research Association, Chicago, Illinois, 1972. (ERIC Number ED 064 894)

McIsaac, D. N. The people of the state: A description through trend surface analysis. 1973. (ERIC Number ED 082 673)

McIsaac, D. N. Trend surface analysis. In H. J. Walberg (Ed.) Evaluating educational performance. Berkeley: McCutchan, 1974.

Smith, N. L. Techniques for the analysis of geographic data in evaluation. Evaluation and Program Planning, 1979, 2, 119-126.

Method: Social Area Analysis/Ecological Analysis

Purpose: Social Area Analysis (Ecological Analysis) is a set of procedures which allow an evaluator to assess the relationships between an intervention program or service delivery system and the demographic characteristics of a particular geographic region.

Why and When to Use: Social Area Analysis would be appropriate for any evaluation where relationships between groups and geographic areas are a major concern, especially evaluations of education, health, and welfare service delivery. Social Area Analysis is not a discrete procedure, but a point of view. It is also a collection of techniques used to study characteristics of subpopulations within defined geographic areas such as census tracts, counties, or service regions called catchment areas. Social Area Analysis employs such familiar techniques as factor analysis, cluster analysis, and multiple regression.

Basic Procedures: 1. Define the geographic region or catchment areas of interest and collect relevant demographic data. 2. Through factor analysis or cluster analysis, develop theoretically meaningful and psychometrically stable indices of catchment area characteristics. 3. Through profile analysis, identify similar catchment areas. 4. Collect data on conditions, behaviors, or characteristics of interest, and study the relationships between these variables and catchment area indices through multiple regression. 5. Verify apparent relationships through direct inquiries of group members, experimentally controlled studies of treatment interventions, or time series designs.

Advantages/Benefits: Social Area Analysis requires no special data collection procedures except the collection of extensive demographic information which is usually available from census records.

Disadvantages/Costs: Social Area Analysis involves considerable data collection and extensive statistical analysis.

Resources Required: Suitable computer programs with access to a computer.

Basic References: General introductions to Social Area Analysis are provided by:

Kay, F. D., Jr. Applications of social area analysis to program planning and evaluation. Evaluation and Program Planning, 1978, 1, 65-78.

Smith, N. L. Techniques for the analysis of geographic data in evaluation. Evaluation and Program Planning, 1979, 2, 119-126.

Struening, E. L. Social areas analysis as a method of evaluation. In E. L. Struening & M. Guttentag (Eds.), Handbook of evaluation research (Vol. 1.). Beverly Hills, CA: SAGE Publications, 1975.

A presentation of the basic procedures is provided by:

Shevky, E., & Bell, W. Social area analysis. Stanford: Stanford University Press, 1955.

The most comprehensive statistical presentation can be found in:

Tryon, R. C., & Bailey, D. E. Cluster analysis. New York: McGraw-Hill, 1970.

Method: Concept Analysis

Purpose: Concept analysis helps the evaluator clarify his thinking about general and abstract questions. Examples are: What is meant by compensatory education? What do we mean by effective reading instruction? What does it mean to say that a student knows something?

Why and When To Use: When the questions in an evaluation contain abstract terms, it is recommended that concept analysis be carried out. Concept analysis is also needed when individuals talk past each other using the same terms, or when different people interpret the same data differently.

Basic Procedures: There are three basic procedures—a generic-type analysis, a differentiation-type analysis, and a conditions-type analysis. A generic-type analysis deals with questions of the form, "What is X?" or "What is meant by X?" e.g., "What is compensatory education?" or "What is meant by equality of educational opportunity?" In other words, "What are the basic features which make compensatory education a form of education and provide the criteria for distinguishing compensatory education from non-compensatory education?" Differentiation-type analysis is called for when the concept question in an evaluation requires making distinctions. For example, the question "What is the difference between teaching and indoctrination?" requires one to make distinctions between different forms of education. In differentiation-type analysis we clarify it and make a concept more useful by pointing to the different basic meanings it has. An example of a conditions-type analysis is: "Under what conditions would it be true to say that somebody knows something. The purpose of a conditions-type analysis is to produce the set of necessary and sufficient conditions for the proper application of a concept to any of its many and varied instances.

Advantages/Benefits: Helps to clarify an evaluator's thinking. Helps to answer concept questions on which depend on the answers to value, fact, and technical questions in an evaluation. Costs nothing except the evaluator's time.

Disadvantages/Costs: Difficult to carry out. Can be time consuming with little benefit to an evaluation.

Resources Required: None.

Basic References:

Caulley, D. N. Concept analysis in evaluation. No. 61 of the Paper and Report Series of the Research on Evaluation Program. Portland, OR: Northwest Regional Educational Laboratory, 1981

Green, T. F. The acitivites of teaching. New York: McGraw-Hill, 1971.

Soltis, J. F. An introduction to the analysis of educational concepts. Reading, MA: Addision-Wesley, Second Edition, 1978.

Wilson, J. Thinking with concepts. London: Cambridge University Press, 1963.

Method: Thematic Matrix Analysis

Purpose: To identify themes which characterize the program being evaluated. The themes help to describe and make the program more understandable to the audience.

Why and When to Use: When the evaluator wants to describe to the audience what goes on in a program and what its activities and processes are like. There are two types of situations in which such representations will be particularly useful. The first is when an interested audience is too distant from the program: members of the audience do not know enough about it and cannot experience it directly; under these conditions they may feel uncomfortable about making judgments about the program or decisions about its future. The second is when participants in the program are too closely involved in its functioning to be able to step back from it and see it as a whole, yet different and unique to each participating observer. In both situations, there is a need for a thematic representation of the program that will allow its audience to experience something of the program, and perhaps to understand it better than before.

Basic Procedures: Interview participants of the program, observe the program, and analyze documents related to the program. Let the themes emerge from the data rather than preordinately beginning with themes. Make a matrix of themes crossed with instances of where they occur in the program.

Advantages/Benefits: Themes help to organize and to make the data meaningful. Themes help to portray the realities of a program.

Disadvantages/Costs: In order for themes to emerge, the evaluator must invest a considerable amount of his/her time in the evaluation.

Resources Needed: None.

Basic References:

Della-Piana, G. M. Literary and film criticism. In N. L. Smith (Ed.), Metaphors for evaluation. Beverly Hills: SAGE Publications, 1981.

Kemmis, S. Telling it like it is: The problem of making a portrayal of an educational program. Urbana, Illinois: Center for Institutional Research and Curriculum Evaluation, University of Illinois, 1974.

Method: Document Analysis

Purpose: Document analysis is the analysis of documents in order to gather facts.

Why and When To Use: Document analysis is superior in finding out retrospective information about a program, and may be the only way that certain information is available. Document analysis is particularly useful at the beginning of an evaluation when the evaluator is trying to understand why the program is the way it is. Documents are an excellent source for determining the purposes, rationale, and history of a program. Doing document analysis is usually a useful prelude to collecting new data. Knowing the purposes of a program helps the evaluator decide what data are to be collected. Doing document analysis one finds out when data have already been collected and what new data need to be collected.

Basic Procedures: Tracking is a process of working through documents looking for information that will confirm some hypothesis. Content Analysis is the creation of categories in order to analyse documents. Case study aggregation is a means for aggregating diverse case studies together using a common conceptual framework so that the findings will be cumulative.

Advantages/Benefits: 1. Document analysis is superior to interviewing for collecting some kinds of retrospective data. 2. Information obtained from documents is often more credible than information obtained via observation and interviewing. 3. Documents are convenient to use. 4. Documents are often available free or at little cost. 5. Documents are non-reactive: That is, it is not usual to find masking or sensitivity because the producer knows he or she is being studied by some social scientist. 6. Records save the time and money that original data collection requires. 7. Program documents provide the evaluator with information about many things that cannot be observed because they may have taken place before the evaluation was begun and because they may include private interchanges to which the evaluator is not directly privy.

Disadvantages/Costs: 1. Often, documents are written to make a program look good and thus can be misleading. 2. There is a dependency on the memory of the person doing the reporting. 3. Although reams of material may be available, it often will not contain much of the needed information or be sufficiently detailed. 4. Documents may reflect clerical lapses, typographical errors, biases, or outright deception. 5. Agency records may be inaccurate, out of date, or months behind on entries. 6. The definitions and categories used by an agency's records may be inappropriate for evaluation purposes. 7. Documentary facts never come "pure", since they are always refracted through the mind of the recorder. 8. Documents may provide unrepresentative samples.

Resources Required: None.

Basic References:

Caulley, D. N. Document analysis in program evaluation. No. 60 of the Paper and Report Series of the Research on Evaluation Program. Portland, OR: Northwest Regional Educational Laboratory, 1981.

Guba, E. G. Investigative reporting. In N. L. Smith (Ed.), Metaphors for evaluation: Sources of new methods. Beverly Hills: SAGE Publications, 1981.

Guba, E. G., and Lincoln, Y. S. Effective evaluation. San Francisco: Jossey-Bass, 1981.

Murphy, J. T. Getting the facts. Santa Monica: Goodyear, 1980.

Method Description Sheet 12

Method: Legislative History

Purpose: A legislative history is constructed to determine the true intent of legislators regarding a law enacted by the legislative branch of a government.

Why and When to Use: Legislative histories can be done with those evaluative studies which are based on programs having a history in the legislature. ESEA Title I, Title IV, and the Title I Technical Assistance Centers are examples of programs which have their direct source in federal legislation. Questions asked include: "Are the aims and objectives of the program the same as those intended by the legislature?" and "Are the outcomes of the program the same as those intended by the legislature?" Such questions are significant when the legislators are an important audience of the evaluation.

Basic Procedures: The essential steps in doing a legislative history are as follows:

1. Identify critical sources of official information.
2. Get the bill, congress, and statute number for the particular subject under examination.
3. Find and examine both House and Senate Committee reports, as well as any conference reports.
4. Check the proceedings of each house and read the debates therein.
5. Examine the hearings and any other allied material.

Advantages/Benefits: By doing a legislative history, the evaluator is able to learn the true intent of legislators regarding a program which has been mandated by the legislature and which is being evaluated. By knowing the legislators' intent the evaluator, can be more responsive to the legislators' information needs. No direct costs are involved.

Disadvantages/Costs: Difficult for the evaluator to find his/her way around legislative documents. Time consuming to carry out. Documents recording legislative intent are not as extensive for the state legislature as for the federal legislature.

Resources Required: Access to a library that has legislative documents. The acquaintence of a government document librarian should be sought to help the evaluator find his/her way around documents.

Basic References:

Caulley, D. N. Legislative history and evaluation. No. 62 of the Paper and Report Series of the Research on Evaluation Program. Portland, OR: Northwest Regional Educational Laboratory, 1981.

Folson, G. B. Legislative history. Charlottesville: University Press of Virginia, 1972.

Method: The Key Interview

Purpose: The key interview has a variety of purposes: to confront the interviewee with the results of the evaluation to that point; to elicit the interviewee's response to allegations about the program being evaluated; to test the evaluative data for accuracy; to get the interviewee's side of the story; and to get new information that is available only from this knowledgeable source.

Why and When to Use: The key interview is appropriate when evaluators find themselves in situations in which a key informant (for instance, a project director) is hostile and uncooperative, as, for example, in a third-party evaluation in which the informant feels threatened.

The key interview is also appropriate when the evaluator wants to test the first draft of an evaluation report for the following: to determine its credibility to various audiences, to solicit their reactions, and to get further details that may help clarify issues being studied.

Basic Procedures: The key interview typically involves at least the following steps.

(1) Gaining entree. Some audiences may refuse to see the evaluator; others may argue passively; still others may actively seek out the evaluator to present their side of the story

(2) Preparing. The evaluator (and at least one other person who will accompany the evaluator) should prepare by rereading the entire accumulated file of material, formatting the interview, and writing and arranging the questions that will be asked.

(3) Beginning the interview. The interest of the evaluators in finding out the facts should be stressed. The questioning should begin with easier questions that are not likely to be emotionally tinged.

(4) Controlling the interview. The evaluator should establish and maintain control of the interview throughout.

(5) Confronting the interviewee. Eventually the interviewee should be confronted with the evaluation findings that have been developed so far.

(6) Recording. Everything that transpires at the interview should be recorded.

Advantages/Benefits: To confirm evaluative facts and interpretations, and to obtain crucial new information.

Disadvantages/Costs: May alienate key audiences to an evaluation.

Resources Required: Two evaluators and a tape recorder.

Basic References:

Guba, E. G. Investigative journalism. In N. L. Smith (Ed.), New techniques for evaluation. Beverly Hills, CA: SAGE Publications, 1981.

Williams, P. N. Investigative reporting and editing. Englewood Cliffs, NJ: Prentice-Hall, 1978.

Method Description Sheet 14

Method: Interviewing: Circling, Shuffling and Filling

Purpose: The terms, "circling," "shuffling," and "filling," are three techniques for extending and rounding out information.

Why and When to Use: When the evaluator has incomplete information on a topic and wants to use interview data to extend and round out that information, then circling, shuffling and filling are appropriate techniques.

Basic Procedures: Circling involves taking information obtained from one source and running it back around the evaluator's circle of contacts for refutation or confirmation. The contacts are assumed to be cooperative.

Shuffling is like circling but with two differences: the contacts are not assumed to be cooperative, and the evaluator makes an effort not only to confirm or deny but also to build upon the original bit of information. For example, when person B is presented with the tidbit of information gleaned from person A and is asked to confirm or deny it, he is also pressed for more details. If he denies the information, he is asked what is really true, what really did happen. If he confirms the information, he is asked to supply more details. Whatever person B supplies (and, of course, whatever person A had supplied) are then presented to person C, and so on round the chain.

Filling involves two things: building up a perimeter around the investigative area--a boundary--and then working to fill in the gaps that exist within it.

Advantages/Benefits: To confirm or disconfirm evaluative facts and interpretations, and to obtain crucial new information

Disadvantages/Costs: Circling, shuffling and filling may be very time consuming.

Resources Required: None.

Basic References:

Guba, E. G. Investigative journalism. In N. L. Smith (Ed.), *New techniques for evaluation*. Beverly Hills: SAGE Publications, 1981.

Williams, P. N. *Investigative reporting and editing*. Englewood Cliffs, NJ: Prentice-Hall, 1978.

Method Description Sheet 15

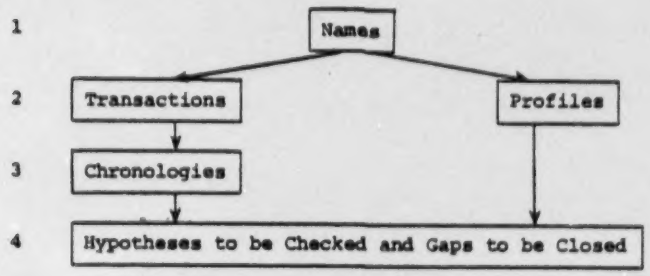
Method: Documentary Files and Summaries

Purpose: Files and summaries are an important permanent record and serve as the basis for the evaluation as the evaluator writes it.

Why and When to Use: Files and summaries have a great deal more utility than simply as a convenience in cataloging collected material and serving as a memory aid. They are, in fact, an important tactical investigatory tool in their own right.

Basic Procedures:

Stage



1. Filing begins with folders for each of the individuals involved in the evaluation (Stage 1).

2. As the persons originally involved enter into transactions with others (a transaction may be as simple as a meeting for lunch), folders are developed related to each transaction (Stage 2).

3. Simultaneously, profiles are developed on each of the persons for whom files have been established.

4. Chronologies are developed for the various transactions (Stage 3).

5. As the files develop, important items are cross-referenced.

6. Each file is also summarized on a systematic basis, so that the evaluator need not deal with the bulky original items each time he enters the file, but rather with a summary of what it contains.

7. Entries in the several files will give rise to hypotheses to be checked or will point to informational gaps that probably ought to be closed (Stage 4).

Advantages/Benefits: Files and summaries are valuable in ordering, and keeping track of, a large amount of information. The approach is especially fruitful in the monitoring of a program.

Disadvantages/Cost: Maintaining and keeping files can be very tedious and timeconsuming.

Resources Required: Facilities for filing.

Basic Reference:

Guba, E. G. Metaphor adaptation report: Investigative journalism, No. 4 of the Paper and Report Series of the Research on Evaluation Program, Portland, OR: Northwest Regional Educational Laboratory, 1978.

Method Description Sheet 16

Method: Cost-Benefit Analysis

Purpose: To compare the costs and benefits to society of various policy alternatives

When and Why to Use: Cost-benefit analysis can be used whenever all the benefits can be readily converted into pecuniary values or when those that cannot be converted tend to be unimportant or can be shown to be similar among the alternatives that are being considered.

Basic Procedures: For each policy alternative, estimate the costs and convert the benefits into pecuniary terms. For each alternative calculate a cost-benefit ratio and select the alternative with the lowest cost-benefit ratio.

Advantages/Benefits: The advantage of cost-benefit analysis is the range of comparisons that can be made among alternatives, not only within a particular service area, but among them. Thus, educational projects can be ranked according to their cost-benefit ratios, but they can also be compared with health and transportation projects that have been evaluated for their costs and benefits.

Disadvantages/Costs: The disadvantage is that benefits and costs must be assessed in pecuniary terms. It is not often possible to do this systematically and rigorously. For example, while the gains in earnings and certain self-provided services attributed to higher levels of literacy (say) might be assessed according to their pecuniary worth, how does one assess such benefits as improvement in self-esteem of the newly-literate adult populations or their enhanced appreciation of reading materials?

Resources Needed: None.

Basic References:

Rothenberg, J. Cost-benefit analysis: A methodological exposition. In M. Guttentag and E. L. Struening (Eds.), Handbook of evaluation research, Vol. 2. Beverly Hills: SAGE Publications, 1975.

Levin, H. M. Casebook on cost analysis in educational evaluation (No. 33 of the Paper and Report Series of the Research on Evaluation Program). Portland, OR: Northwest Regional Educational Laboratory, 1979.

Method: Cost-Effectiveness Analysis

Purpose: To compare the costs and benefits to society of various policy alternatives

When and Why to Use: There are many evaluative situations in which it is possible to assess the educational results of particular alternatives, while not being able to place pecuniary values on those results. For example, it is difficult to conceive of acceptable methods for placing pecuniary worth on test score gains in arithmetic operations or on improvements in self confidence. In these cases, it seems more appropriate to compare the results of programs in terms of their effectiveness on some criterion or set of criteria, then comparing these measures of effectivenes with the costs of the alternative programs.

Basic Procedures: For each policy alternative, estimate the costs and measure the effectiveness in terms common to all alternatives. For each alternative, calculate a cost-effectiveness ratio and select the alternative with the lowest ratio.

Advantages/Benefits: It is commonly found that the most "effective" approach is not always the most cost-effective. Yet, without an analysis of costs it will not be possible to know this. Further, the adoption of the most "effective" alternative can actually cost many times as much as the most cost-effective one. Cost-effectiveness comparisons require only that the impact of alternative strategies, along with their respective costs, be derived, while cost-benefit analysis requires that we put a monetary value on the impact. The results of policy-oriented experiments or quasi-experiments lend themselves naturally to cost-effectiveness comparisons.

Disadvantages/Costs: Cost-effectiveness is of no value for selecting among alternatives that address different objectives. Measuring the effectiveness of programs requires the setting of experimental or quasi-experimental designs, which is difficult and time consuming.

Resources Required: None.

Basic References

Levin, H. Cost-effectiveness analysis in evaluation research. In M. Guttentag and E. L. Struening, (Eds.), Handbook of evaluation research, Vol. 2. Beverly Hills: SAGE Publications, 1975.

Levin, H. M. Casebook on cost analysis in educational evaluation (No. 33 of the Paper and Report Series of the Research on Evaluation Program). Portland, OR: Northwest Regional Educational Laboratory, 1979.

Method Description Sheet 18

Method: Cost-Utility Analysis

Purpose: To compare the costs and benefits to society of various policy alternatives

When and Why to Use: When benefits cannot be measured in monetary terms or in terms of effectiveness, cost utility is the alternative. When subjective assessment must be made about the nature and probability of educational benefits as well as their relative worth, cost utility analysis may be an appropriate tool.

Basic Procedures: First, the decision maker uses the information available to him or her to ascertain the probability of achieving particular educational outcomes with each of the policy alternatives. With the cost-utility approach there can be more than one outcome per alternative. Second, the decision maker places relative values on each of the educational outcomes to weight them according to their desirability. The method for doing this is to rate each potential outcome on a scale of utility which reflects the desirability of that outcome. For example, a decision maker could assess the value of each outcome on a 0-10 point scale with equal intervals, in which 10 represents the highest value. Third, the subjective probability of each outcome is multiplied by the utility placed upon the outcome and adding these products across outcomes. This calculation is done separately for each policy alternative. Finally, costs are divided by the expected utilities to obtain cost utility ratios for each alternative.

Advantages/Benefits: Cost-utility analysis permits the use of a wide range of types of both qualitative and quantitative data in forming the decision. The advantages of the cost-utility approach are that the data requirements are less stringent than cost-benefit or cost-effectiveness approaches, that a large number of potential outcomes can be included in the evaluation, and that imperfect information and uncertainty can be addressed systematically.

Disadvantages/Costs: The highly subjective nature of the assessments of effectiveness and the values placed upon them by the decision maker prevents the kind of replicability from analysis-to-analysis that might be obtained with the more stringent cost-benefit and cost-effectiveness approaches.

Resources Required: None.

Basic References:

Edwards, W., Guttentag, M., & Snapper, K. A decision-theoretic approach to evaluation research. In E. L. Struening and M. Guttentag, (Eds.), Handbook of evaluation research, Vol. 1. Beverly Hills: SAGE Publications, 1975.

Stokey, E., & Zeckhauser, R. A primer for policy analysis. New York: W. W. Norton & Co., 1978.

Levin, H. M. Casebook on cost analysis in educational evaluation (No. 33 of the Paper and Report Series of the Research on Evaluation Program). Portland, OR: Northwest Regional Educational Laboratory, 1979.

Method: Cost-Feasibility Analysis

Purpose: To compare the costs to society of various policy alternatives

When and Why to Use: Cost-feasibility analysis represents a method of estimating the costs of an alternative in order to ascertain whether it can be considered. That is, if the cost of any alternative exceeds the budget and other available resources, there is no point in doing any further analysis. As a concrete illustration, one might view the situation of compensatory education in which a specified amount is available for augmenting the education of each disadvantaged child. If this amount is $400 per child, then any alternative that exceeds this constraint would not be feasible.

Basic Procedures: The cost of every policy alternative is determined.

Advantages/Benefits: Cost-feasibility analysis can determine whether it is necessary to do the following: estimate the benefits of a cost-benefit analysis, determine the effectiveness of a cost-effectiveness analysis, or estimate the utilities of a cost-utility analysis.

Disadvantages/Costs: Cost-feasibility represents a limited form of analysis which can only determine whether alternatives are within the boundaries of consideration. It cannot be used to determine which ones should actually be selected.

Resources Required: None.

Basic References:

Levin, H. M. Casebook on cost analysis in educational evaluation (No. 33 of the Paper and Report Series of the Research on Evaluation Program). Portland, OR: Northwest Regional Educational Laboratory, 1979.

Method Description Sheet 20

Method: Blanket Sampling

Purpose: To cover as much of the entire span of events as is possible

Why and When to Use: Blanket sampling is used when the sampler wants to cover a wide range of events.

Basic Procedures: The sampler moves about freely, and fairly constantly, from location to location so as to cover a wide range of events.

Advantages/Benefits: This method of sampling does not constrain the sampler to follow a given session or event through to completion.

Disadvantages/Costs: If photography is being used to blanket sample, the photographer will gather a rather large proportion of liminal events, generally at the expense of sequences showing development of action, changes of groups, shifts in proxemics, etc. Basically, this is the methodology of traditional photojournalism, and the fact that it tends to gather only liminal events probably accounts for both the strong impact and the uncontrolled biases of most photographic documentaries.

Resources Required: None, except if using photography; then camera equipment will be needed by the evaluator.

Basic References:

Templin, P. A. Photography as an evaluation technique (No. 32 in the Paper and Report Series of the Research on Evaluation Program). Portland, OR: Northwest Regional Educational Laboratory, 1979.

Templin, P. A. Handbook in evaluating with photography (No. 63 in the Paper and Report Series of the Research on Evaluation Program). Portland, OR: Northwest Regional Educational Laboratory, 1981.

Method Description Sheet 21

Method: Shadow Sampling

Purpose: To carry out detailed observations of how an individual or individuals spend their time

Why and When to Use: Shadow sampling is used when it is desired to find out what activities the participants of a program engage in.

Basic Procedures: Shadow sampling involves following a single subject through a program and recording the experiences and interaction of that individual.

Advantages/Benefits: The data collected will all share a common independent variable: the individual being shadowed.

Disadvantages/Costs: The individual chosen for shadowing may introduce systematic bias into the data through the types of interactions and situations in which she or he typically and habitually engages (or does not engage). Taking this viewpoint, shadow sampling may result in collecting data about the individual, rather than about the program. This problem can be overcome, to some extent, by shadowing several subjects, either simultaneously or consecutively. However, simultaneous shadowing generally requires multiple observers, which is likely to be very obtrusive, while consecutively shadowing several subjects through a program of any length may require unrealistic amounts of contact time.

Resources Required: One or more observers and time to carry out the observations

Basic References:

Templin, P. A. Photography as an evaluation technique (No. 32 in the Paper and Report Series of the Research on Evaluation Program). Portland, OR: Northwest Regional Educational Laboratory, 1979.

Templin, P. A. Handbook in evaluating with photography (No. 63 in the Paper and Report Series of the Research on Evaluation Program). Portland, OR: Northwest Regional Educational Laboratory, 1981.

Method Description Sheet 22

Method: Time-Based Sampling

Purpose: Time-based sampling reveals shifting patterns of persons in a fixed space over time.

Why and When to Use: Time-based sampling can be extremely useful in studies of the use of space and can also provide information on the changing relations between groups and between members of groups.

Basic Procedures: The spatial coverage (e.g., a room) is fixed and samples are taken at regular intervals of time. One method of collecting data with this method is to use a 16mm movie camera fitted with an extreme wide-angle lens (often covering a whole room), and an intervalometer. Alternatively, a 35mm still camera can be used, fitted with a bulk film magazine, motor drive, extreme wide-angle lens, and intervalometer. The advantage of the movie camera is that the resulting film can be projected as a time-lapse movie, thus showing motion patterns and changes in the spatial distribution of persons in the environment. On the other hand, time-based photo sequences taken with the 35mm still camera can be examined minutely for such details as eye contact, and prints can be measured to quantify the proximity of subjects to one another.

Advantages/Benefits: Time-based sampling is easy to set up and carry out.

Disadvantages/Costs: Collecting data on a constant time basis may tend to exaggerate the importance of the large proportion of "inactive" time that tends to separate human events.

Resources Required: None, except if using photography, then camera equipment will be needed by the evaluator.

Basic References:

Templin, P. A. Photography as an evaluation technique (No. 32 in the Paper and Report Series of the Research on Evaluation Program). Portland, OR: Northwest Regional Educational Laboratory, 1979.

Templin, P. A. Handbook in evaluating with photography (No. 63 in the Paper and Report Series of the Research on Evaluation Program). Portland, OR: Northwest Regional Educational Laboratory, 1981.

Method Description Sheet 23

Method: Event-Based Sampling

Purpose: To sample as many events of a certain type as is possible

Why and When to Use: Event-based sampling is used when the research question posed for the study is concerned not with a program or system as a whole, but with a particular type of action, interaction, sequence, or event within the overall context. Events may be relatively broad ("question-and-answer periods," "mealtime groupings"), or very specific ("handshakes between persons of the opposite sex after the presentation").

Basic Procedures: The sampler selects a particular type of event and samples the occurrence of it.

Advantages/Benefits: Event-based sampling is easy to set up and carry out.

Disadvantages/Costs: Event-based sampling is generally only appropriate in response to specific research questions. It should probably be accompanied by some more global technique (e.g., blanketing the event area, having another person take field notes, tape recording) to provide contextual information for interpreting the category of specific events.

Resources Required: None, except if using photography; then camera equipment will be needed by the evaluator.

Basic References:

Templin, P. A. Photography as an evaluation technique (No. 32 in the Paper and Report Series of the Research on Evaluation Program). Portland, OR: Northwest Regional Educational Laboratory, 1979.

Templin, P. A. Handbook in evaluating with photography (No. 63 in the Paper and Report Series of the Research on Evaluation Program). Portland, OR: Northwest Regional Educational Laboratory, 1981.

Method Description Sheet 24

Method: Dimensionally Based Sampling

Purpose: To collect data from the widest possible range of types of situations

Why and When to Use: Dimensionally based sampling would appear to be useful where the initial aim of the study is descriptive appraisal, but the constraints of time, space, and work force make it impossible to cover all events in the program.

Basic Procedures: Any information available about the program being evaluated (field notes, previous reports, informant's reports, printed programs, informational brochures, etc.), is examined to determine a set of apparent dimensions for classifying the events and components of the program. For example, in evaluating a conference, the evaluator used the dimensions of subject matter of sessions and meeting format. The subject matter of sessions was divided into "theoretical" versus "applied." The meeting formats were categorized into: a single long presentation, a series of topically related presentations, panel discussions, workshops, etc. The categories of subject matter were cross-indexed with the categories of meeting format, and sampling was carried out across the resultant cells. Events are then selected to represent the largest possible range of dimensions and combination of dimensions. Other constraints, such as observing a whole event, or spending a predetermined amount of time at each event, may also be imposed.

Advantages/Benefits: The advantage of the sampling procedure is that it offers some guarantee of representativeness in the description of the program being evaluated.

Disadvantages/Costs: Dimensions are chosen at the beginning of the evaluation but will usually need to be progressively modified during the course of the study to reflect the specific research questions that arise as a result of the analysis of the data.

Resources Required: None, except if using photography, then camera equipment will be needed by the evaluator.

Basic References:

Templin, P. A. Photography as an evaluation technique (No. 32 in the Paper and Report Series of the Research on Evaluation Program). Portland, OR: Northwest Regional Educational Laboratory, 1979.

Templin, P. A. Handbook in evaluating with photography (No. 63 in the Paper and Report Series of the Research on Evaluation Program). Portland, OR: Northwest Regional Educational Laboratory, 1981.

Method: Research Briefs

Purpose: A research brief is a written and condensed statement that gives important information about research.

Why and When to Use: A research brief is used when part of the expected audience to an evaluation report will not read the complete report but will read a brief outline.

Basic Procedures: Report summary: One of the best known research briefs is probably the summary that appears at the end of most, if not all, final research reports.

Executive summary: A variation of the report summary, the executive summary is placed at the beginning of the report.

Memos: One simple application is the attachment of a single-page memo to the top of the research report as a means of capturing audience attention and introducing the report.

Googles: A google is a one-liner designed to make its owner appear intelligent and well read. An ever moderate supply of googles should enable one to speak intelligently about any of several evaluation studies.

Embedded quotations: This is either an abridgment or a direct restatement of a portion of a research report and is set apart within the text of the report.

News items: The news item is similar to the report summary but appears in newspapers or newletters.

Abstracts: This is a very concise, accurate and informative report summary.

Advantages/Benefits: The research brief helps the reader to grasp the essentials of an evaluation report without the necessity of investing the time and effort reading the whole report.

Disadvantages/Costs: The audience of the brief could be misled through incomplete information or insufficient qualification.

Resources Required: None.

Basic Reference:

Macy, D.J. Research briefs. In N.L. Smith (Ed.), Communication strategies in evaluation. Beverly Hills, CA: SAGE Publications, 1982.

Method Description Sheet 26

Method: Appeals Procedures

Purpose: To review evaluation findings to determine whether they adequately support the claims made

Why and When to Use: After completing a draft evaluation report, the evaluator can share the draft with the program staff to verify the factual accuracy of the report's content and interpretations before finalizing the report For large summative evaluations, it would be prudent to have the procedures reviewed systematically by an independent party, and an independent judgment could be made as to whether the conclusions reached by the evaluator are warranted by the supporting data.

Basic Procedures: Set up an appeals process by which the content and findings of an evaluation report may be questioned by audiences to the evaluations.

Advantages/Benefits: If audiences feel they have been misrepresented in an evaluation reports or if inappropriate instruments or data analysis procedures were used, the audience can have recourse through an appeals process.

Disadvantages/Costs: The appeals procedures may allow the program staff to delete legitimate references to program weaknesses or substantiated interpretations.

Resources Required: None.

Basic References:

Green, M. Legal procedure. *Collier's encyclopedia*. Chicago: Collier Publishing, 1965, *14*, 438-453.

Owens, T. R. & Owen, T. R., Law. N. L. Smith (Ed.), *Metaphors for evaluation*. Beverly Hills: SAGE Publications, 1981.

Sobeloff, S., & Reitze, C. *Standards relating to criminal appeals: American Bar Association Project on Minimum Standards for Criminal Justice*. Tentative Draft. New York: Institute of Judicial Administration, 1969.

Method Description Sheet 27

Method: Storytelling

Purpose: To create a believable and coherent story from the myriad educational events and activities that take place in an educational setting during a specified period of time

Why and When to Use: Creating a story of an evaluation program occurs not only at the end of the evaluation inquiry when the evaluation report is written, but also from the first moments of the evaluation effort. The emphasizing of particular facts and events represents a fundamental structuring of information that is basic to storytelling.

Basic Procedures: A structure or plot is devised to enable the evaluator to choose what information should be collected, how it should be collected, and from whom it should be collected. Grapple with the question of purpose. Why is the story (evaluation) being told? To what end will the story (evaluation) be used? Choose what information should be collected. Identify those essential components, the major events and existents, that are necessary for the development of the plot. Decide from whom information should be collected and how it should be collected.

Focus the story (evaluation) by relating all of the events and existents that were investigated during the conduct of the evaluation (story) to each other and to the overall purpose of the inquiry. This will enable the evaluator to monitor whether they "made sense" in relation to each other and whether they add to the unity and coherence of the investigation as a whole.

Advantages/Benefits: Devising a plot helps the evaluator to choose what information should be collected, how it should be collected, and from whom it should be collected.

Disadvantages/Costs: Creating a plot at the beginning of an evaluation does not allow new facts and issues to emerge as the evaluation progresses.

Resources Required: None.

Basic References:

Wachtman, E. L. Evaluating and storytelling: The narrative quality and structure of educational evaluation. In N. L. Smith (Ed.), Communicating in evaluation: Alternative forms of representation. Beverly Hills: SAGE Publications, 1982.

Method Description Sheet 28

Method: Compelling the Eye

Purpose: To compose (design) an evaluation so that the mind of the decision maker is compelled to deal with the minor decisions, but in a way that major attention is always turned back to the major decision.

Why and When to Use: A request for an evaluation is typically an indication that a major decision is to be made. Just as typically, a number of minor decisions must be made. This major and minor decision composite is analogous to the artist's major point of interest and minor interest points.

Basic Procedures: As the artist is concerned with compelling the eye, the evaluator is concerned with "compelling the mind." If the evaluation is designed so that it attends only to the major decision, the evaluator has created a quick exit avenue from the evaluation problem: thus it becomes an evaluation flavored by its failure to deal with the total evaluation problem. If the evaluation is designed so that too much energy is expended on any or all of the minor decisions, or if it doesn't compel the decision maker to return to the major decision, it is again a flavored evaluation.

Advantages/Benefits: There is a balanced approach in attending to major and minor decisions.

Disadvantages/Costs: The writer of the evaluation report must be skillful in order to produce a report that "compels the mind."

Resources Required: None.

Basic References:

Gephart, W. J. Watercolor painting. In N. L. Smith (Ed.), Metaphors for evaluation. Beverly Hills: SAGE Publications, 1981.

Kautzky, T. The Ted Kautzky pencil book. New York: Van Nostrand Reinhold, 1979.

Method: Representation of Reality

Purpose: To write evaluation reports that represent the reality of the programs being evaluated.

Why and When to Use: The technique can be used for reports of evaluations of programs when it is desired to give the reader a vivid representation of reality.

Basic Procedures: The device used to represent reality is the use of a "stream of consciousness" or "criterion monologue." In representing reality we are given not just one person whose consciousness is tendered but many persons, with frequent shifts from one to the other. Reality (time passed) is violated to get at reality (the conscious wanderings of the mind). The treatment of time is the characteristic and distinctive feature of the technique. Reality is represented by the interior monologues of the characters of the story. The time it takes for such interior monologues is certainly, in reality, not as long as the time it takes to read or hear them. The exterior reality of the moment is nothing but an occasion to release things seen only by reflection in consciousness and not tied to the present or the occurrence which releases them. There are three characteristics of this representation of reality.

(1) Most objective facts appear as reflections in the consciousness of several characters.
(2) The reality of time passed in exterior events is violated by devoting more time to get at the reality of conscious wanderings of the mind.
(3) The things that happen to a few individuals in a random moment represented in completeness, exploited fully, is favored because it reveals more of the elementary things our lives have in common than the course of a human life over a prolonged time, arbitrarily pruned and isolated.

Advantages/Benefits: The technique gives an added dimension of understanding to the reader of an evaluation report.

Disadvantages/Costs: It takes considerable writing skill to portray reality.

Resources Required: None.

Basic References:

Della-Piana, G. M. Literary and film criticism. In N. L. Smith (Ed.), Metaphors for evaluation. Beverly Hills: SAGE Publications, 1981

Auerbach, E. Memesis: The representation of reality in Western literature. Princeton: Princeton University Press, 1953.

Method Description Sheet 30

Method: Accurate, Sharp Descriptions

Purpose: To write evaluation reports that involve accurate, sharp, loving descriptions.

Why and When to Use: This technique can be used for evaluation reports of programs that are regarded by the evaluator as meritorious in some way.

Basic Procedures: Write up the evaluation report as an accurate, sharp, loving description.

Advantages/Benefits: Accurate, sharp, loving description of a program makes the program come alive for the reader. The reader gains a vivid understanding of how the program operates.

Disadvantages/Costs: It takes considerable writing skill to write an accurate, sharp, loving description.

Resources Required: None.

Basic References:

Della-Piana, G. M. Literary and film criticism. In N. L. Smith (Ed.), Metaphors for evaluation. Beverly Hills: SAGE Publications, 1981.

Jarrell, R. The third book of criticism. New York: Farrar, Straus, Giroux, 1965.

Method Description Sheet 31

Method: Graphic Displays

Purpose: To get the people in the target group to respond to a visual image which will communicate a message in a manner the group can readily understand.

Why and When to Use: Graphic displays can be used to clarify complex data, to emphasize relative magnitudes, to help identify trends, and to demonstrate relationships.

Basic Procedures: The first step toward using the techniques of graphic display to improve the reporting of evaluation results is to incorporate them appropriately into the body of our traditional printed evaluation reports in the form of well-designed illustrations, graphs, charts and so on. The next step is to supplement these written reports by other forms of presentation according to the needs of the audiences and as the availability of resources permit. Examples of ways to present the results of evaluations that lend themselves readily to graphic display techniques are exhibits, posters, brochures, overhead tranparencies, slide/tape presentations, videotapes, and films.

Advantages/Benefits: Graphic displays help the audience understand the evaluative information that we present.

Disadvantages/Costs: One difficulty with the use of graphic techniques is that they require more time and money for both report planning and execution.

Resources Required: Access to a graphic designer.

Basic Reference:

Hathaway, W.E. Graphic display procedures. In N.L. Smith (Ed.), Communication strategies in evaluation. Beverly Hills, CA: SAGE Publications, 1982.

Method Description Sheet 32

Method: The Stem and Leaf Displays and Box Plots

Purpose: To display a batch of data so that its properties are visually discernible.

Why and When to Use: The stem and leaf display provides a visual display of the shape of a distribution similar to that provided by a histogram. Box plots are used when the evaluator wishes to explore the properties of his/her data and when he/she wants to give the reader of the evaluation report a visual presentation of the data. This exploration is necessary as a basis for decisions about the questions for which answers might reasonably be sought and the types of analysis which might most appropriately be used. Through such exploration, more formal models of the phenomena can be developed and their goodness of fit tested with both the original and subsequent data. Traditional statistical analysis is designed more for the latter task of model fitting than for the former task of data exploration. Indeed, a confirmatory data analysis restricts the analyst's attention to the model and limits the potential conclusions about the model to "fitting" or "not fitting." Confirmatory procedures give almost no guidance on how to explore data.

Basic Procedures: For the stem and leaf display the data are grouped into categories. The beginning digits of each category become part of the stem. Each category starts a new row. The last digits of the scores of each category are displayed on the row. These last digits are referred to as the leaves. A box plot is a visual representation of the central tendency and the spread of a set of scores. The box locates the middle 50% of the data, extending from the 25th percentile to the 75th percentile. The location of the median within the box gives some indication of the symmetry of the data. Dotted lines extend outward from the box to the location of the scores closest to, but inside, what are termed the "two inner fences." The fences are boundaries beyond which the occurrence of scores could be considered unusual. The inner fence contains 99% of the distribution. Scores outside the inner fences are highlighted individually as large dots in the display. The box plot thus highlights unusual scores.

Advantages/Benefits: The stem and leaf display has two advantages over the histogram. One is that it retains all the original data so that the content of the groupings within the batch remain apparent, making detailed examination easier. The other is that it is easier to prepare than a histogram so that no graph need be drawn. The box plot gives ready visual inspection of the location and shape of the distribution of the data. The box plot provides a summary of the data which is resistant to the influence of deviant data and yet does not obscure them. The box plot as a simple graphical display helps the evaluator to look at the data in more than one way. It can suggest further analyses. The box plot can also be very helpful in the evaluator's attempts to communicate with others, particularly with lay audiences.

Disadvantages/Costs. Lay audiences cannot interpret the stem and leaf display or the box plot unless they are given prior tutoring in the meaning of the stem and leaf display or the box plot.

Resources Required: None.

Basic References:

McGaw, B. Exploratory data analysis. In N. L. Smith (Ed.), New techniques for evaluation. Beverly Hills: SAGE Publications, 1981.

Tukey, J. W. Exploratory data analysis. Reading, MA: Addison-Wesley, 1977.

Method Description Sheet 33

Method: Still Photography

Purpose: To describe evaluative information by means of photographs.

Why and When to Use: The following is a listing of conditions for which evaluators might choose to use photography:

- a record of events in detail.
- visual information that is of primary interest.
- tracking the activities of a single participant in the program.
- thick description of the process in a program rather than outcomes.
- a record of information about social interaction of particular individuals, groups, or communities.
- documenting regular routines of regular and special classes or programs.
- activities rather than goals of a program.
- understanding the subjective nature of the participants' experience.

Basic Procedures: The procedures for the data collection are as follows:

- Taking pictures.
- Classifying them, focusing the topic, selecting the most relevant.
- Gathering verbatim data by interviewing and observing.
- Photointerviewing or showing photographs to the subjects for their interpretation.
- Going back frequently to confirm or disconfirm the data, staying in the study site as long as possible.
- Repeating the process.
- Selecting and interpreting photographs, and arranging a final report.

Advantages/Benefits: Photography constitutes a form of visual communication that provides relevant nonverbal information and basic imagery not available by other means.

Disadvantages/Costs: A photograph is open to a wide range of subjective interpretation.

Resources Required: Photographic equipment.

Basic References:

Templin, P.A. Photography as an evaluation technique. No. 32 of the Paper and Report Series of the Research on Evaluation Program, Portland, OR: Northwest Regional Educational Laboratory, 1978.

Templin, P.A. Still photography in evaluation. In N.L. Smith (Ed.)., Communication strategies in evaluation. Beverly Hills, CA: SAGE Publications, 1982.

Method Description Sheet 34

Method: Oral Briefings

Purpose: An oral briefing is a personal interaction between evaluator and audience in which the results of evaluation are presented orally. The purpose is to increase the impact of the evaluation on the audience.

Why and When to Use: Oral briefings are used when the evaluator wants to have more impact on the audience than just a written report would have.

Basic Procedures:

(1) Planning the briefing. A successful briefing must be carefully planned far in advance of the actual briefing date.

(2) Setting the stage. The first way is to determine the participants, both among the presenters and the audience. A second way to set the stage for the briefing is to provide advance briefing materials to each audience member.

(3) Presenting the briefing. At the beginning the presenter captures the audience's attention. Then the presenter must provide three basic types of information: (a) a description of the program, (b) a comparison of this information with appropriate other information to provide perspective, and (c) any interpretations of the findings and their implications.

(4) Following up. Often a briefing can crystallize a number of followup actions. First, the assignments need to be defined clearly and completely. Second, responsibility for each action needs to be assigned to a specific individual. Third, staff must be assigned the responsibility of tracking the assigned task and its completion. Fourth, it is necessary for the formal assignment to include a specific role for the evaluator.

Advantages/Benefits: Briefings, being visible events involving several persons, possess an impact that is difficult to ignore. Briefings do not have the impersonal exchange of a written report.

Disadvantages/Costs: Briefings are transitory and not lasting, and thus are open to misinterpretation or misrepresentation by audience members. The presentation is almost entirely dependent on the personal abilities of the single presenter. The nature of a briefing requires omitting from the presentation much useful information.

Resources Required: None.

Basic Reference:

Hendricks, M. Oral policy briefings. In N.L. Smith (Ed.), Communication strategies in evaluation. Beverly Hills, CA: SAGE Publications, 1982.

Method: Briefing Panel Presentations

Purpose: A briefing panel is a group of people, perhaps a project's advisory committee, called together to hear a compressed presentation of project activities and findings. Its members raise questions, discuss issues, and perhaps offer additional interpretations and recommendations. The burden of the evaluators is to devise more pungent presentations, anticipate data needs, and incorporate reactions.

Why and When to Use: Evaluators can use a briefing panel to learn how better to inform others through their evaluation reports.

Basic Procedures: An example of a briefing panel presentation is as follows: Panelists who have some expertise in the area being evaluated were chosen. The panelists were charged: (1) to consider specifically the strengths and shortcomings of the program strategy and to consider more generally the various strategies and obstacles to such improvement; and (2) to discuss the philosophy, methods, and results to date of the program being evaluated.

Proceedings began at 9 a.m. with an hour set aside for presentations by evaluators and witnesses. At 10 a.m. the panelists were asked to break an imposed silence and to indicate questions or topics on which they would like more information, which they got. At 11 a.m. the panelists were asked to take charge and to proceed in any way they chose. They were asked to spend at least an hour on their task. The issues raised and the alternative ideas proposed were included in the interim evaluation report.

Advantages/Benefits: External panel experts can apply a fresh perspective in interactions with program-related personnel, can provide multidisciplinary expertise in project reviews, and can permit use of nationally known experts who would be too expensive to hire on an extended basis.

Disadvantages/Costs: The reaction of most participants was that the proceeding would have been improved with more structure. Panel members and some observers indicated they felt confused about what the panel was supposed to do and what other persons were to do.

Resources Required: None.

Basic Reference:

Stake, R.E., & Balk, D.E. Briefing panel presentations. In N.L. Smith (Ed.), Communication strategies in evaluation. Beverly Hills, CA: SAGE Publications, 1982.

Method Description Sheet 36

Method: Adversary Hearings

Purpose: To adjudicate between opposing issues of a program which is being evaluated.

Why and When to Use: The adversarial model operates with the assumption that truth emerges from a hard, but fair fight, in which opposing sides, after agreeing upon the issues in contention, present evidence in support of each side. The fight is refereed by a neutral figure, and all of the relevant evidence is weighed by a neutral person (or persons) to arrive at a fair result.

Basic Procedures:

The Legislative Phase: In the legal system, a dispute may arise only when there is a rule of law to provide the standards against which an act or series of events may be evaluated. The analogy of law in program evaluation is a statement of the goals and objectives of a program. Wolf (1975) used an elaborate dragnet to garner complaints in a phase he called "issue generation," followed by an issue selection phase.

Retrieval Process: Wolf (1975) turned over the task of issue delimitation to two teams of evaluators, one of which presented the case for the existing program, and the other the case for alternatives.

The Trial of Hearing: A trial is held in which alternative sides of an issue are presented.

The Jury or the Panel: Wolf (1975) selected highly qualified professionals representing a variety of viewpoints on the program in question. Alternatively, a jury can be selected.

Advantages/Benefits: Adversary approaches provide a disciplined method for introducing and testing qualitative testimony. Cross-examination provides a form of control to reveal the limits of any bit of testimony to support a proposition.

Disadvantages/Costs: The adversarial process, as it has been used so far, is time-consuming, expensive, and cumbersome.

Resources Required: Time, money and participants.

Basic References:

Levine, M. Scientific method and the adversary model. Some preliminary thoughts. American Psychologist, 1976, 29, 661-677.

Levine, M. Adversary hearings. In N.L. Smith (Ed.), Communication strategies in evaluation. Beverly Hills, CA: SAGE Publications, 1982.

Wolf, R. Trial by jury. A new evaluation method. I. The Process. Phi Delta Kappan, 1975, 57, 185-187.

Method Description Sheet 37

Method: Committee Hearings

Purpose: Committee hearings have been used for two purposes. One purpose is to evaluate an evaluation (i.e. meta-evaluation). A hearing is held to examine the evaluative procedures, to test the conclusions reached, and to examine the recommendations. Another purpose is valuation, i.e., assignment of merit or worth through the evaluative process.

Why and When to Use: Use when there are contending issues that need to be resolved surrounding an evaluation.

Basic Procedures:

Committee Selection: It is quite possible that an existing committee can serve in a committee hearing. School boards, textbook committees, curriculum committees, advisory panels, faculty committees, and executive committees may be used. Membership on committees could also be determined by identifying what kind of expertise would provide an adequate examination of data treatment and interpretations.

Roles: Roles that can be considered are the committee chairperson, committee counsel, and witnesses.

Stages of Operation: The committee membership should have a role in preparing guidelines and in selecting witnesses or identifying the type of witnesses they would like to make presentations at the hearing. Witnesses should be informed about the operation of the hearings. Prior to the hearings, the committee should review the procedures to be used. The actual format of the hearing will vary with the purpose of the hearing. The final phase of the committee work is that of summary.

Advantages/Benefits: The committee hearing ensures that an evaluation will be used, not set aside to gather dust.

Disadvantages/Costs: The committee hearing can be time-consuming, expensive and cumbersome.

Resources Required: Time, money and participants.

Basic Reference:

Stenzel, N. Committee hearings. In N.L. Smith (Ed.), *Communication strategies in evaluation.* Beverly Hills, CA: SAGE Publications, 1982.

Method Description Sheet 38

Method: Television Presentations of Hearings

Purpose: To provide high-quality information to those who are making decisions and implementing programs. Television presentations have a major purpose, that of providing information and clarifying issues—not of judging or of rating individual programs.

Why and When to Use: Hearings are designed to serve an educational function by providing a public forum for discussion of a controversial topic from different and often competing perspectives. Many decision makers cannot attend hearings, so videotapes of the proceedings are circulated to decisionmakers. Television seems best suited to evaluation studies that have a naturalistic mode of inquiry. If the purpose of the evaluation is to tell a story, to describe a dynamic, complex program in operation, to trace patterns of decisionmakers, or to ascertain people's attitudes and opinions, television seems an ideal medium for reporting evaluation results.

Basic Procedures: Television is particularly suited to evaluations that end with an event such as a courtroom hearing. The hearing is not intended to result in a victory for one side or the other. There is no jury present to enter a final judgment, and all decisions are left to the viewing audience. As with all technology, it is best to involve professionals who know how to produce and direct educational television programs.

Advantages/Benefits: Television can:

- record an event with very little distortion.
- be used to reach large and scattered audiences.
- transport the viewer to a scene (classroom) or gain access to an individual (politician, celebrity) not normally accessible to the viewer.
- make complex ideas and abstractions more concrete through the use of examples or demonstrations.

Disadvantages/Costs:

- Personal face-to-face interaction between viewer and communicator are eliminated.
- Reality can be misrepresented through selective reporting and editing.
- Costs of production are usually much higher than those for other media, particularly print media.
- Complex phenomena that requires long verbal explanations, statistical data, or specialized knowledge, may be difficult to portray visually.

Resources Required: Access to television equipment and personnel. Costs may range into the hundreds of thousands of dollars.

Basic Reference:

Shoemaker, J.S. Television presentations. In N.L. Smith (Ed.), *Communication strategies in evaluation.* Beverly Hills, CA: SAGE Publications, 1982.